Wine Quality Prediction

Machine Learning

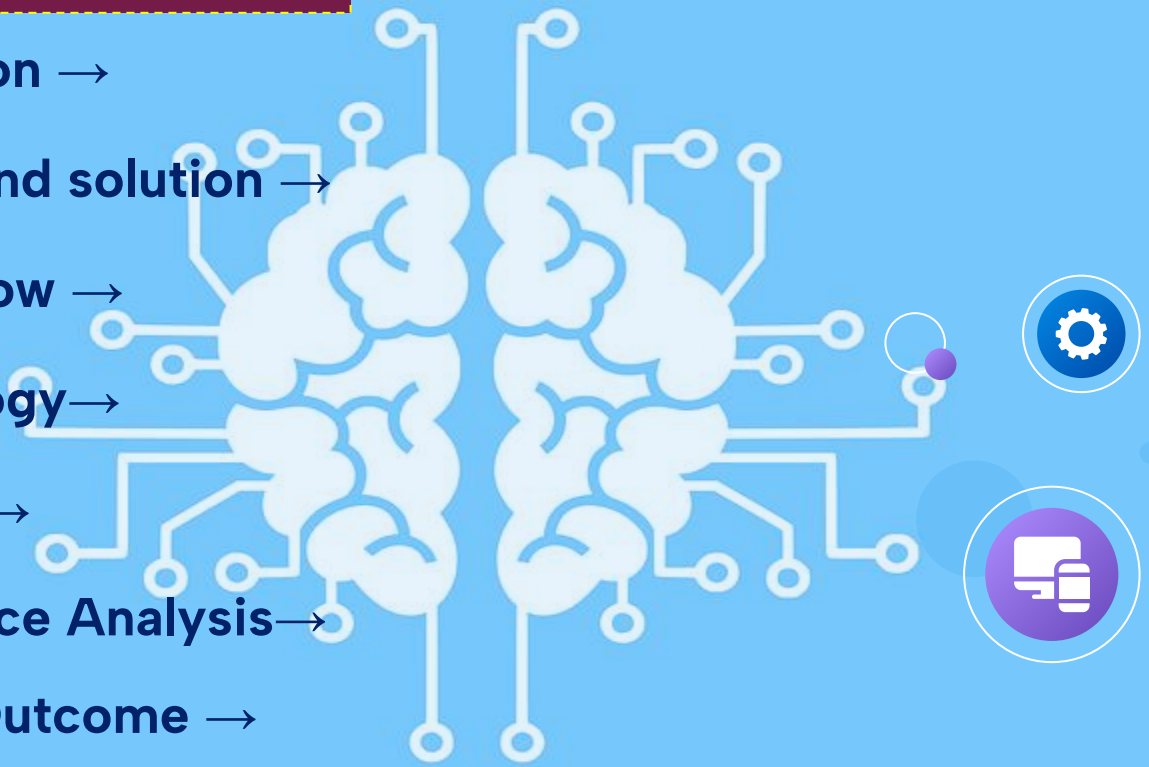
VIVA

Group – antidote

1. Sikta Sonawane (36107)
2. Sonal Diwan (36108)
3. Sonam Gutte (36109)

Diems , Sambhaji Nagar.

Table of contents



01

Introduction →

History

- The wine industry has a long-standing tradition. However, the process of assessing and predicting wine quality is a complex task that requires expertise and experience.
- With the advent of machine learning techniques, it has become possible to develop models that can effectively predict wine quality based on various factors and attributes.

Introduction

relying solely

leveraging machine learning

wine quality prediction

gained significant attention

recent years

analyze large volumes

machine learning algorithms

gain valuable insights

algorithms

objective measurements

quality

detect patterns

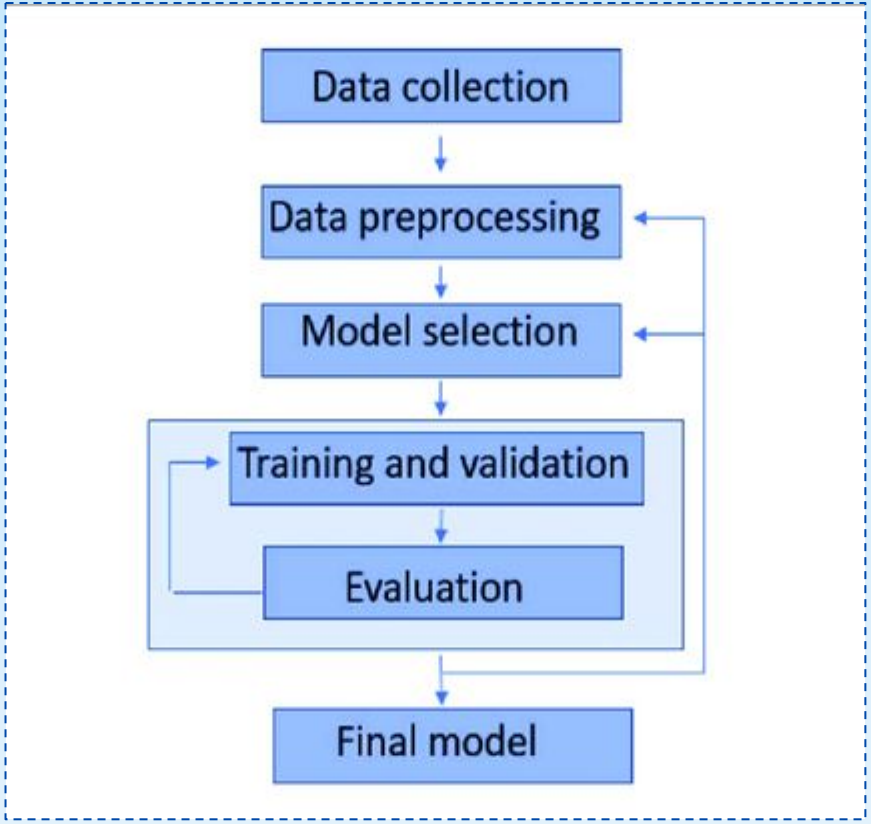
wine based

make accurate predictions

Problem Statement →

"To develop an accurate and reliable machine learning model that can predict the quality of wines based on a set of input variables, aiming to provide winemakers and enthusiasts with an objective and efficient method for assessing and predicting wine quality."

02. System Flow



03

Methodology→

Methodologies

1. Feature Selection

This may involve statistical techniques, domain knowledge, or feature importance ranking algorithms on this basis we can choose most relevant features

2. ML Algorithms

The selection of algorithms depends on the nature of the problem, the dataset, and the desired model performance. Commonly regression model is used.

3. Model training

This involves splitting the data into training and testing sets, defining appropriate evaluation metrics, and optimizing the model's parameters through techniques like grid search or cross-validation.

4. Evaluation & deployment

Assessment of the trained model's performance & deploying the trained model into a production environment, which may involve developing an interface or application


Random Forest
Logical Regression
~ Data set rows count - 6463

04

Platforms→

Flask

web development,
one drop at a time

matplotlib

scikit
learn

05

Result & Outcome →

127.0.0.1:5000/predict

Wine Quality Prediction

pH:

Alcohol:

fixed acidity:

volatile:

citric acid:

residual sugar:

chlorides:

free sulfur dioxide:

total sulfur dioxide:

density:

sulphates:

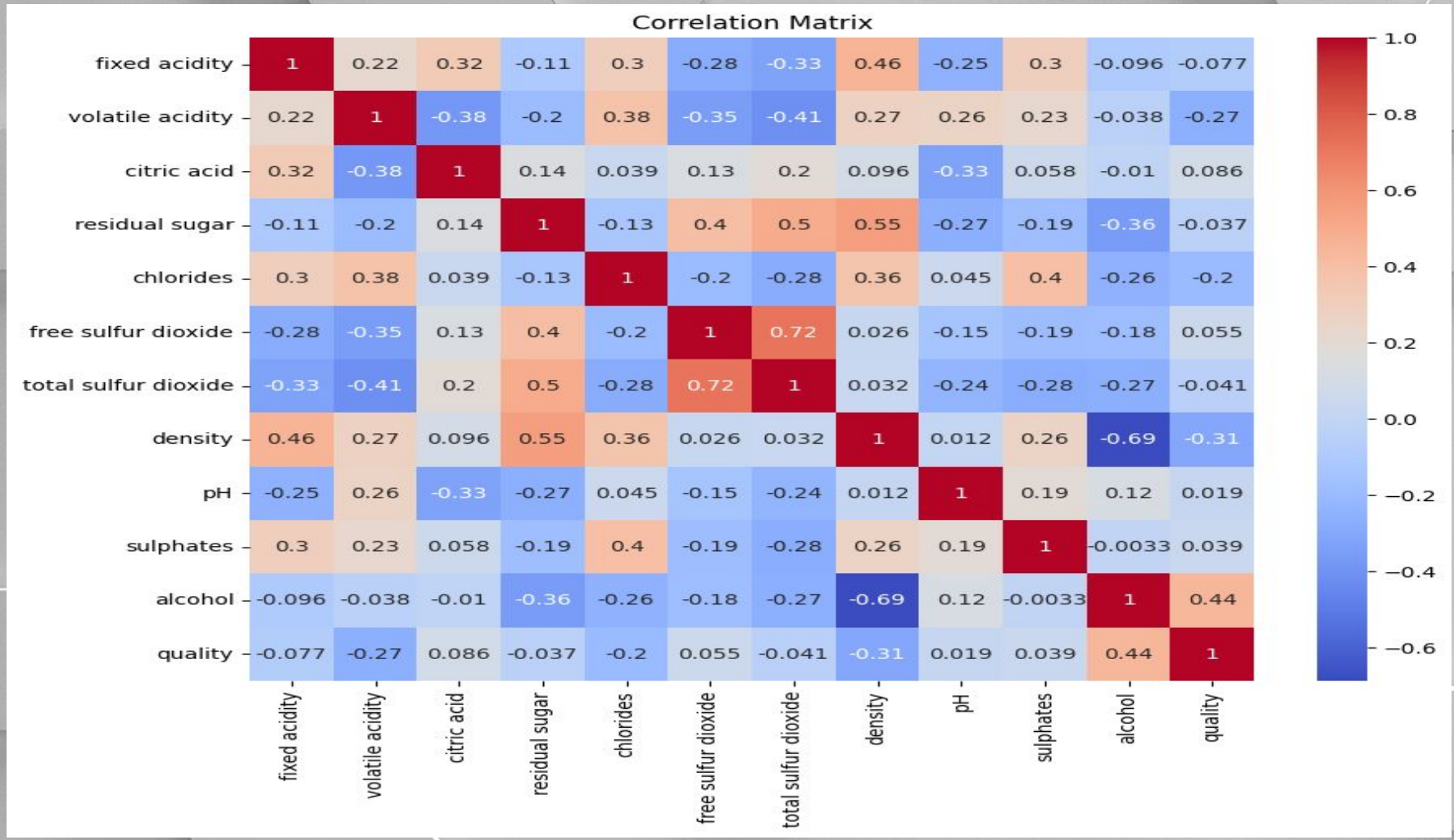
Predict

Predicted Wine Quality: 4.65


06
Performance
Analysis→

Role of Attributes in performance

1. *Fixed Acidity:* Higher levels of fixed acidity
2. *Citric Acid:* natural acid found in fruits
3. Residual Sugar: balanced level of residual sugar
4. *Chlorides*: A moderate level
5. *Sulphates:* Sulphates, when present in appropriate amounts, can act as antioxidants and contribute to the wine's stability.
6. *Alcohol:* Generally, good quality wines tend to have an appropriate and well-balanced

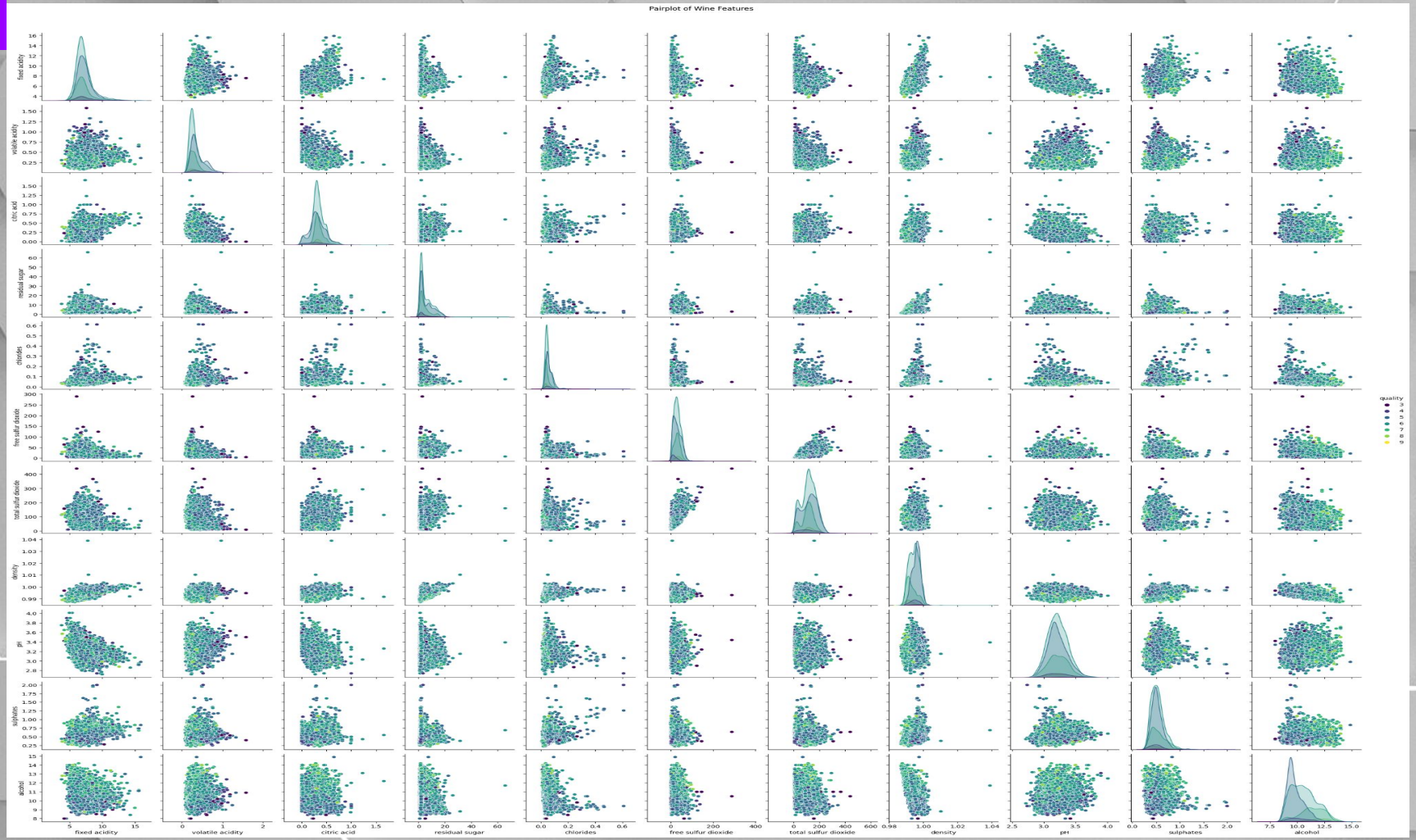

Correlation Matrix
fixed acidity 1 0.22 0.32 -0.11 0.3 -0.28 -0.33 0.46 -0.25 0.3 -0.096 -0.077
volatile acidity 0.22 1 -0.38 -0.2 0.38 -0.35 -0.41 0.27 0.26 0.23 -0.038 -0.27
citric acid 0.32 -0.38 1 0.14 0.039 0.13 0.2 0.096 -0.33 0.058 -0.01 0.086
residual sugar -0.11 -0.2 0.14 1 -0.13 0.4 0.5 0.55 -0.27 -0.19 -0.36 -0.037
chlorides 0.3 0.38 0.039 -0.13 1 -0.2 -0.28 0.36 0.045 0.4 -0.26 -0.2
free sulfur dioxide -0.28 -0.35 0.13 0.4 -0.2 1 0.72 0.026 -0.15 -0.19 -0.18 0.055
total sulfur dioxide -0.33 -0.41 0.2 0.5 -0.28 0.72 1 0.032 -0.24 -0.28 -0.27 -0.041
density 0.46 0.27 0.096 0.55 0.36 0.026 0.032 1 0.012 0.26 -0.69 -0.31
pH -0.25 0.26 -0.33 -0.27 0.045 -0.15 -0.24 0.012 1 0.19 0.12 0.019
sulphates 0.3 0.23 0.058 -0.19 0.4 -0.19 -0.28 0.26 0.19 1 -0.0033 0.039
alcohol -0.096 -0.038 -0.01 -0.36 -0.26 -0.18 -0.27 -0.69 0.12 -0.0033 1 0.44
quality -0.077 -0.27 0.086 -0.037 -0.2 0.055 -0.041 -0.31 0.019 0.039 0.44 1
fixed acidity
volatile acidity
citric acid
residual sugar
chlorides
free sulfur dioxide
total sulfur dioxide
density
pH
sulphates
alcohol
quality
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6

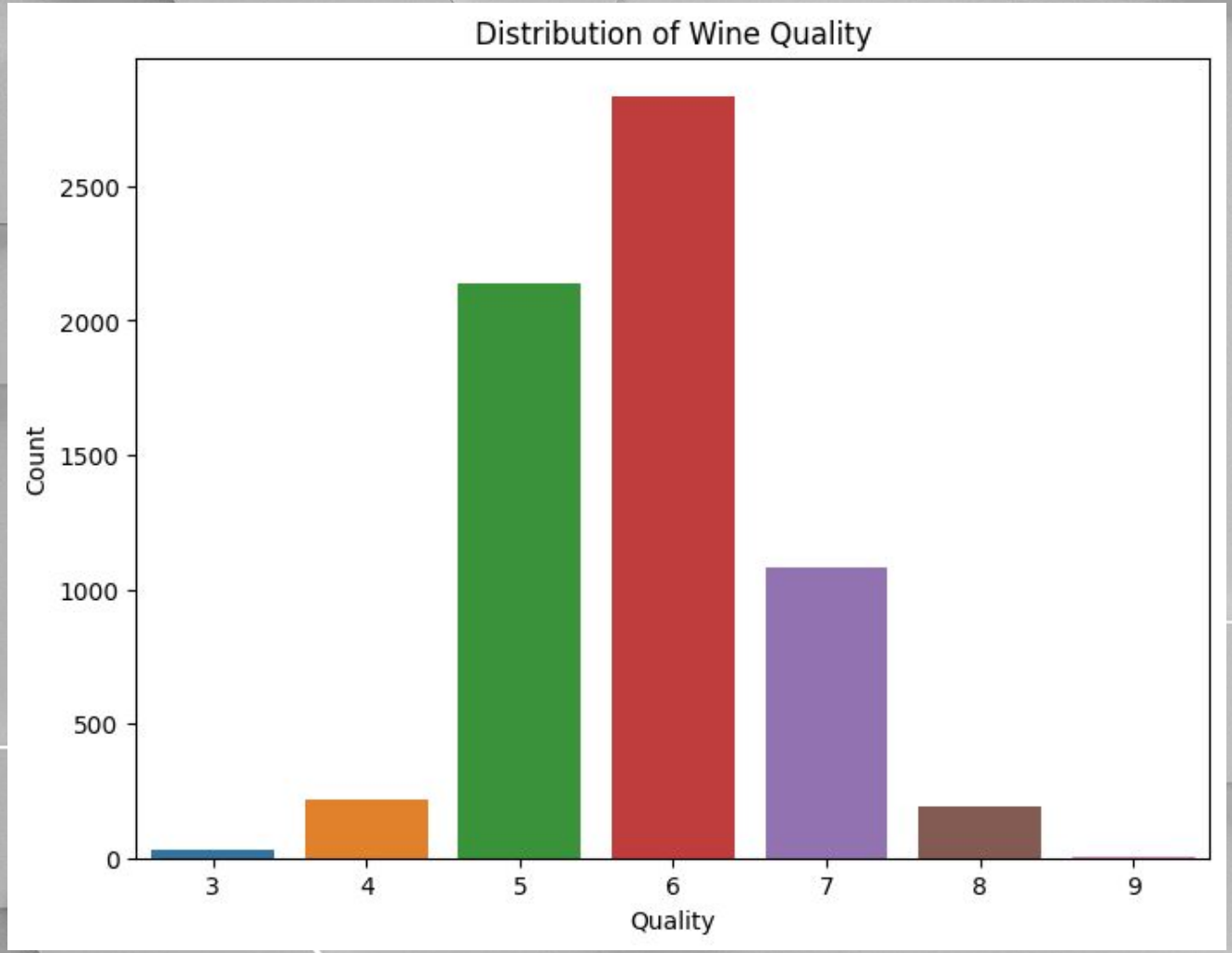
Accuracy & Confusion matrix

```
Accuracy: 1.00
Classification Report:
              precision    recall  f1-score   support

         red       0.99      0.99      0.99       322
       white       1.00      1.00      1.00       971

    accuracy                           1.00      1293
   macro avg       0.99      0.99      0.99      1293
weighted avg       1.00      1.00      1.00      1293

Confusion Matrix:
[[319   3]
 [  3 968]]
```

Performing better on the validation data with less difference between the validation and training data.

Pair plot


Pairplot of Wine Features
fixed acidity
volatile acidity
citric acid
residual sugar
chlorides
free sulfur dioxide
total sulfur dioxide
density
pH
sulphates
alcohol
quality
3
4
5
6
7
8
9


Distribution of Wine Quality
Count
2500
2000
1500
1000
500
0
3
4
5
6
7
8
9
Quality

07 Conclusion

1. The project's findings and the developed model have several practical implications in the wine industry.
2. Winemakers can utilize the model for quality control, production optimization, and grape selection, leading to consistent quality standards and improved production processes.
3. Wine sellers and distributors can benefit from the model's insights in evaluating wines, determining pricing strategies, and offering personalized recommendations to consumers.
4. Furthermore, the project contributes to the broader field of machine learning applications in the food and beverage industry, showcasing the potential for data-driven insights in traditional domains.


Thank You !